

CLAIMS

What is claimed is:

1. A method of forming information for determining a direction of an acoustic source using at least three spaced-apart microphones, the microphones coupling acoustic signals from at least two pairs of microphones with each pair of microphones receiving two acoustic signals and having a separation distance and an orientation of its two microphones, the method comprising:

for each pair of microphones, calculating a plurality of sample elements

for the two acoustic signals received by the pair of microphones, the plurality of sample elements corresponding to a ranking of possible time delays between the two acoustic signals received by the pair of microphones with each sample element having a time delay and a numeric sample value;

for the plurality of sample elements of each pair of microphones, mapping each sample element to a sub-surface of potential acoustic source locations according to its time delay and the orientation and the separation distance of the pair of microphones for which the sample element was calculated, and assigning the sub-surface the sample value of the sample element, producing a plurality of sub-surfaces for each pair of microphones;

for a boundary surface intersecting each of the plurality of sub-surfaces,
the boundary surface divisible into a plurality of cells, calculating
a weighted value in each cell of the boundary surface by
combining the sample values of the plurality of sub-surfaces
proximal the cell to form a weighted surface with the weighted
value of each cell of the weighted surface being indicative of the
likelihood that the acoustic source lies in a direction of a bearing
vector passing through the cell.

2. The method of claim 1, further comprising:

calculating a likely direction to the acoustic source by determining the
bearing vector to the cell of the weighted surface having a
maximum magnitude.

3. The method of claim 2, further comprising:

storing the likely direction as metadata of an audio-visual event
associated with the generation of the acoustic signals.

4. The method of claim 1, further comprising: storing the weighted
values.

5. The method of claim 1, wherein the plurality of sample elements of each pair of microphones is calculated by cross-correlating the acoustic signals received by the pair of microphones during a time window.

6. The method of claim 5, further comprising:
pre-filtering the acoustic signals prior to cross-correlation.

7. The method of claim 5, wherein cross-correlating is performed using a generalized cross-correlation function.

8. The method of claim 1, wherein each sub-surface of potential acoustic source locations is a hyperboloid.

9. The method of claim 1, wherein each sub-surface of potential acoustic source locations is a cone.

10. The method of claim 9, wherein calculating the weighted value in each cell further comprises:

for each pair of microphones, interpolating the sample values between neighboring sub-surfaces on each cell of the boundary surface to form for each pair of microphones an acoustic location function having a resampled value on each cell; and

in each cell, combining the resampled values of each of the acoustic location functions.

11. The method of claim 10, wherein in each cell the resampled values are combined by summing the resampled values of each of the acoustic location functions on the cell.

12. The method of claim 1, wherein there are three or more pairs of microphones.

13. The method of claim 1, wherein there are four microphones and two pairs of microphones.

14. The method of claim 1, wherein there are four microphones and six pairs of microphones.

15. The method of claim 1, wherein the boundary surface is a hemisphere.

16. The method of claim 2, wherein the likely direction is used to select a camera view of the acoustic source.

17. The method of claim 2, wherein the likely direction is used to control a camera view of the acoustic source.

18. The method of claim 2, wherein the likely direction is stored as metadata for a visual recording of the acoustic source.

19. A method of forming information for determining the location of an acoustic source using at least three spaced-apart microphones, the microphones coupling acoustic signals from at least two pairs of microphones with each pair of microphones receiving two acoustic signals and having a separation distance and an orientation of its microphones, the method comprising:

for each pair of microphones, cross-correlating the two acoustic signals received by the pair of microphones to produce a plurality of sample elements with each sample element having a time delay and a sample value;

for each sample element of the plurality of sample elements associated with each pair of microphones, mapping the sample element to a cone of potential acoustic source locations appropriate for the time delay of the sample element and the separation distance and the orientation of the pair of microphones for which the sample element was calculated and assigning the cone the sample value of the sample element, forming a sequence of cones for each pair of microphones;

for each pair of microphones, mapping the sequence of cones associated with the pair of microphones to a boundary surface divisible into

a plurality of cells and interpolating the sample values between adjacent cones to form a continuous acoustic location function on the boundary surface having a resampled value in each cell, thereby forming a plurality of acoustic location functions; and in each cell, combining the resampled value of each of the acoustic location functions to form a weighted acoustic location function having a weighted value in each cell indicative of the likelihood that the acoustic source lies in a direction of a bearing vector passing through the cell.

20. The method of claim 19, further comprising:
pre-filtering the signals prior to cross-correlation.

21. The method of claim 20, wherein the pre-filtering is performed using a phase transform filter.

22. The method of claim 19, wherein the resampled values are combined in each cell by summing the resampled values on the cell.

23. The method of claim 22, wherein the boundary surface is a hemisphere.

24. The method of claim 23, wherein there are four microphones arranged as a rectangular array with one microphone disposed on each corner of a rectangle and the hemisphere has an origin coincident with the center of the rectangle.

25. The method of claim 24, wherein the pairs of microphones are two pairs of microphones with each of the two pairs of microphones having a midpoint coincident with the origin of the hemisphere.

26. The method of claim 25, further comprising: at least one additional pair of microphones having a midpoint non-coincident with the origin of the hemisphere.

27. The method of claim 26, wherein there are four non-coincident pairs of microphones.

28. The method of claim 19, further comprising:
temporally smoothing the weighted acoustic location function of one time window with the weighted acoustic location function of at least one previous time window.

29. The method of claim 19, wherein a sample rate and the separation distance between the two microphones of each pair of microphones is selected so

that the number of sample elements for each pair of microphones is greater than 90° divided by a desired cell resolution in degrees.

30. The method of Claim 29, wherein the number of sample elements is greater than 180° divided by a desired cell resolution in degrees.

31. A method of forming information for determining the location of an acoustic source using at least three spaced-apart microphones, the microphones coupling signals from at least two pairs of microphones with each pair of microphones receiving two acoustic signals and having a separation distance and an orientation of its microphones, the method comprising:

for each pair of microphones, cross-correlating the two acoustic signals received by the pair of microphones to produce a sequence of discrete sample elements for the pair of microphones with each sample element having a time delay and a sample value;

for each pair of microphones, mapping each sample element of its sequence of sample elements to a cone of potential acoustic source locations appropriate for the time delay of the sample element and the orientation and separation distance of the pair of microphones for which the sample element was calculated, and assigning the cone the sample value, thereby forming for each pair of microphones a sequence of cones;

for each pair of microphones, mapping its sequence of cones to a hemisphere divisible into a plurality of cells and interpolating sample values between adjacent cones to form for each pair of microphones an acoustic location function having a resampled value on each cell of the hemisphere; and

forming a weighted acoustic location function having a weighted value in each cell by combining in each cell the resampled values of each of the acoustic location functions, the weighted value of each cell being indicative of the likelihood that the acoustic source lies in a direction of a bearing vector passing through the cell.

32. The method of Claim 31, wherein a sample rate and a separation between microphones of each pair of microphones is selected so that the number of sample elements for each microphone pair is greater than ninety degrees divided by a desired cell resolution in degrees.

33. The method of Claim 31, further comprising:
selecting a cell having a maximum value; and
calculating the bearing direction from an origin of the microphones that extends in a direction through the cell having the maximum value.

34. The method of Claim 31, further comprising:

temporally smoothing the combined acoustic location function of a current time window with a result from at least one previous time window.

35. A system for generating data regarding the location of an acoustic source, comprising:

at least three microphones coupled to provide acoustic signals from at least two pairs of microphones with each pair of microphones consisting of two microphones receiving two acoustic signals and having a separation distance and an orientation;

an analog-to-digital converter adapted to sample the acoustic signals at a preselected rate and to convert the acoustic signals into digital representations of the acoustic signals;

a correlation module receiving the digital representations of the acoustic signals and outputting for each pair of microphones a sequence of discrete sample elements with each sample element having a time delay and a sample value; and

an acoustic source direction module receiving the sample elements configured to form a weighted acoustic location function on a boundary surface, the acoustic source direction module comprising:

a mapping sub-module mapping each sample element to a cone of potential acoustic source locations appropriate for the time

delay of the sample element and the separation distance and the orientation of the pair of microphones for which the sample element was calculated and assigning each cone the sample value;

a resampling sub-module adapted to interpolate the sample values between adjacent cones of each pair of microphones on the boundary surface, the resampling module forming an acoustic location function for each pair of microphones that has a resampled value on each cell of the boundary surface; and

a combining sub-module configured to combine the resampled values of the acoustic location function on each cell into a weighted value for the cell that is indicative of the likelihood that the acoustic source lies along in the direction of a bearing vector passing through the cell.

36. The system of Claim 35, further comprising:

a speech detection module configured to limit directional analysis to acoustic sources that are human speakers.

37. The system of Claim 35, further comprising:

at least one camera;

a video storage module for storing video data from the at least one camera; and

an offline storage module for receiving and storing acoustic source direction data from the acoustic source direction module.

38. The system of claim 35 wherein the mapping sub-module, resampling sub-module, and combining sub-module comprise program code residing on a memory of a computer.

39. A system for generating data regarding the location of an acoustic source, comprising:

a plurality of pairs of microphones;

correlation means for producing for each pair of microphones a sequence of discrete sample elements with each sample element having a time delay and a sample value; and

acoustic source direction means receiving the sample elements and calculating a weighted value on each of a plurality of cells of a common boundary surface, the weighted value on each cell being indicative of the likelihood that the acoustic source lies in a bearing direction passing through the cell.

40. A computer program product for forming information for determining a direction to an acoustic source from the acoustic signals of at least

three microphones coupled to provide acoustic signals from at least two pairs of microphones with each pair of microphones consisting of two microphones receiving two acoustic signals and having a separation distance and an orientation, the computer program product comprising:

- a computer readable medium;

- a cross-correlation module stored on the computer readable medium, and configured to receive a digital representation of the acoustic signals and outputting for each pair of microphones a sequence of sample elements with each sample element having a time delay and a sample value; and

- an acoustic source direction module stored on the computer readable medium, and configured to receive the sample elements and perform the steps of:

- for the plurality of sample elements of each pair of microphones, mapping each sample element to a sub-surface of potential acoustic source locations according to its time delay and the orientation and the separation distance of the two microphones of the pair of microphones for which the sample element was calculated, and assigning to the sub-surface the numeric sample value of the sample element, producing a plurality of sub-surfaces for each pair of microphones; and

calculating for a boundary surface intersecting each of the plurality of sub-surfaces and divisible into a plurality of cells, a weighted value in each cell of the boundary surface by combining the values of the plurality of sub-surfaces proximal the cell to form a weighted surface with the weighted value of each cell of the weighted surface being indicative of the likelihood that the acoustic source lies in a direction of a bearing vector passing through the cell.